A LOW-POWER SMART VISION SYSTEM-ON-A-CHIP DESIGN FOR ULTRA-FAST MACHINE VISION APPLICATIONS

Wai-Chi Fang

NASA's Center of Integrated Space Microsystems
Jet Propulsion Laboratory, California Institute of Technology
MS 198-219, 4800 Oak Grove Drive, Pasadena, CA 91109-8099, USA
TEL: (818) 354-4695, FAX: (818) 393-5007 , EMAIL: wai-chi.fang@jpl.nasa.gov

Abstract

In this paper, an ultra-fast smart vision system-on-a-chip design is proposed to provide effective solutions for real time machine vision applications by taking advantages of recent advances in integrated sensing/processing designs, electronic neural networks, advanced microprocessors and sub-micron VLSI technology. The smart vision system mimics what is inherent in biological vision systems. It is programmable to perform vision processing in all levels such as image acquisition, image fusion, image analysis, and scene interpretation. A system-on-a-chip implementation of this smart vision system is shown to be feasible by integrating the whole system into a 3-cm x 3-cm chip design in a 0.18–μm CMOS technology. The system achieves one tera-operation-per-second computing power that is a two order-of-magnitude increase over the state-of-the-art microcomputer and DSP chips. Its high performance is due to massively parallel computing structures, high data throughput rates, fast learning capabilities, and advanced VLSI system-on-a-chip implementation. This highly integrated smart vision system can be used for various NASA scientific missions and other military, industrial or commercial vision applications.

KEYWORDS
VISION SYSTEM, ACTIVE PIXEL SENSOR, NEURAL PROCESSOR, SYSTEM-ON-A-CHIP, SMART SENSOR

1. Introduction

Machine vision is a challenging problem [1,2]. The challenge of this problem is caused by the following tasks: variations in the sensing environment, limitations of the sensors, signal-to-noise issues, difficulties of processing algorithms, performance of computation power, real time processing requirements and field deployable requirements. The success of a versatile machine vision system depends on succeeding at all these tasks. In general, the goal of a versatile machine vision system is of determining what and how useful information should be extracted from images and delivering appropriate outputs for various tasks. Many paradigms and algorithms have been proposed over the pass three decades toward this problem. However, a versatile machine vision system has not emerged yet.

Recent advances in sensors, processors, integrated sensing/processing technology, electronic neural networks, and VLSI technology appear to be very promising to provide solutions for real time machine vision problems [3,4,5]. In this paper, an ultra-fast smart vision system-on-a-chip design is proposed to provide effective solutions for real time machine vision applications by taking advantages of advances in integrated sensing/processing designs, electronic neural networks, advanced microprocessors and sub-micron VLSI technology. The proposed smart vision system mimics what is inherent in biological vision systems. Moreover, this vision system is programmable and capable of performing ultra-fast machine vision processing in all levels such as image acquisition, image fusion, image analysis, scene interpretation, and control functions. The system provides about one tera-operation-per-second computing power which is a two order-of-magnitude increase over that of state-of-the-art microcomputers. Its high performance is due to massively parallel computing structures, high data throughput rates, fast learning capabilities, and advanced VLSI system-on-a-chip implementation. This highly integrated smart vision system can be used for various NASA scientific missions and other military, industrial or commercial vision applications.

Section two presents the smart vision algorithm and architecture. Section three describes the system and its on-chip building blocks. Section four describes the on-chip APS camera. Section five describes on-chip neural computer. Section six describes the on-chip microcomputer. Section seven gives a brief conclusion and future work.

2. Smart Vision Processing Model and Algorithm

The human visual system provides a demonstration of a versatile and powerful vision system. For a machine vision system, it is not necessary to copy all the details of the human visual system, but it is important to understand the true complexity behind its power and flexibility. In general, much of the current machine vision research concerns not only understanding the process of vision but also designing effective vision systems for various real-world applications.

The proposed versatile machine vision system is inspired by the human visual system. Its computation model is based on a simplified model of the human visual system as shown in Figure 2.1. This simplified model of a human visual system consists of the following stages: 1. raw image collection, 2. synthetic image generation, 3. images fusion, 4. fused images analysis, 5. semantic interpretation [1]. The ultimate design goal is to build an eye-brain machine which can automatically recognize, localize, and classify point, area and volume objects and phenomena in real-time. In general, the eye-brain machine (EBM) includes two major subsystems: the EBM Eye and the EBM Brain. The EBM Eye is a compact optoelectronic subsystem which integrates a wide range of different sensors with geometric, radiometric, and spectral parameters meeting the actual science and mission requirements. The EBM Brain is a high performance control and data handling subsystem, which provides computing resources to perform various on-board vision tasks.

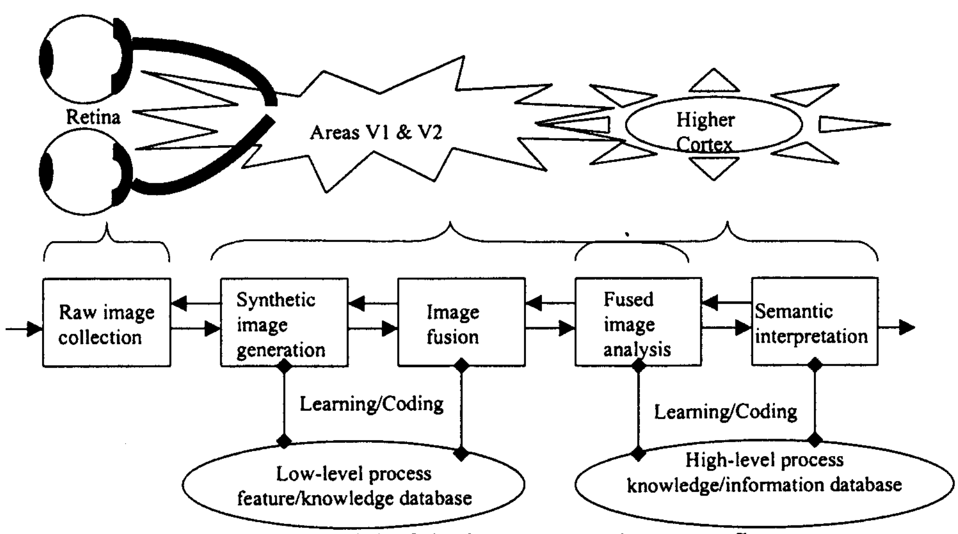

Figure 2.1. A simplified model of the human visual process flow.

3. Smart Vision Processing Architecture and Design

3.1. A Rapid Prototyping of the Eye-Brain Machine by Using Commercial-off-the-shelf Components

Figure 3.1 shows an example system design of the eye-brain machine for the VIGILANTE. The VIGILANTE has been under development at JPL [4]. The VIGILANTE is an ultra-fast smart sensor for target recognition and precision tracking in a simulated cruise missile defense scenario. The VIGILANTE consists the VIGIL (Viewing Imager/Gimbaled Instrumentation Laboratory) and the ANTE (Analog Neural Three-dimensional processing Experiment). VIGIL is an integrated optical system that splits/transmits the incoming light (steered by a gimbaled mirror) detected by the respective IR/visible/UV sensors. ANTE is the processing system that selects each sensor channel for processing.

A rapid prototyping of the VIGILANTE system has been performed based on the system diagram shown in Figure 3.1 by using commercial-off-the-shelf components. An APS camera captures the image and sends it to an image frame grabber via a serial bus. The image frame grabber holds the image and feeds columns or rows of all pre-sequenced 64x64 subwindow to the Column Loading Formatter via the PCI bus in the background. The Column Loading Formatter feeds a column or row of these pre-sequenced subwindows to the Column Loading Input Chip every 250 ns (4 MHz). The 3-DANN-M then produces 64 inner-products (each with one 4096-dimensional image vector and a 4096-dimensional template vector) every 250 ns. The 64 analog outputs of the 3-DANN-M are buffered and converted to 8-bit digital data by using an off-chip ADC array at a rate of 4 MHz. The system thus accomplishes 64 convolutions of a 256x256 image with 64x64 templates in 16 ms at a computation speed of about 1-tera operations per second. These 64 convoluted images generated by the 3DANN-M are stored in the Activity Memory Banks and then passed along to the Post Operation Processor (POP) for desired data fusion and various post operations. Currently, the Activity Memory Banks and the POP are implemented in four Adaptive Solution's CNAPS array processor boards (each PCI-bus based board containing 128 SIMD processors and 32 megabytes of memory). The P6 computer performs command and control of VIGILANTE operations such as image acquisition, templates loading, point operation functions, scene interpretation, data recording, detection/tracking/classification/calibration mode command, etc.

Today's sub-micron fabrication enables designers to put millions of circuits into a single microchip to realize a complete system. Along with this capability, the system design and verification become extremely complex. However, lessons learned from the existing COTS-based prototype and thorough system-level simulations help us to define an effective system-on-a-chip smart vision system. This system level integration approach can eliminate these workaround-type PCI cards used in the VIGILANTE and also alleviate the data bandwidth bottleneck due to the PCI bus.

3.2. A Low-Power Ultra-Fast Eye-Brain Machine by Using System-on-a-chip Technology

Systems-on-a-chip (SOAC) technology enables a system level integration to put an APS camera together with a neural computer and a microcomputer into a single chip. Figure 3.2 shows a system diagram of the proposed smart vision system. An on-chip row/column-parallel image flow architecture is used to connect all on-chip systems and eliminate data bandwidth bottlenecks due to conventional bus architectures. The functional blocks include: (a) an on-chip APS camera, (b) an on-chip programmable neural computer, and (c) an on-chip microcomputer. The operation of each on-chip system is briefly described below.

The on-chip APS camera is used as the optical sensing array in the system. Windowed image data are fed to the on-chip neural neural processor under the control of a smart window handler. The on-chip neural computer is programmed to perform various early vision tasks in high speed. This tera-operation-per-second neural processor serves as a supercomputing engine for various vision processing tasks due to its massively parallel computing structures and its programmability. The on-chip microcomputer is used to perform command and control of the system operation and scene interpretation. The vision system can work with a remote host system through the muti-bus interface unit. The host computer will display the output image or vision science data. A system-on-a-chip implementation of this smart vision system is shown to be feasible by integrating the whole system into a 3-cm x 3-cm chip design in a 0.18–μm CMOS technology (see Table 3.1).

Table 3.1: A smart vision system in a 3-cm x 3-cm chip design in a 0.18–μm CMOS technology

Smart Vision Processor	Technology: 0.18-micron CMOS		
On-Chip System	Building Block	Size (mm sq.)	Power (mW)
On-Chip Camera			
	APS Sensor (1Kx1K)	12.96	100
	Image Frame Memory(1MB)	15.55	100
	Smart Window Handler	6	50
On-Chip Neural Computer			
	Neural Processor (64x64x64)	144	800
	Post Operation Processor	6	200
	Synapse Memory (1MB)	15.55	100
	Activity Memory (1 MB)	15.55	100
			0
On-Chip Microcomputer			0
	PowerPC750 Processor	10	300
	128MB Flash Memory	497.66	200
	512 KB SRAM	7.78	200
	Bus Interface	5.29	200
	FPGA (256Kgate)	6	200
	EEPROM(256KB)	3.89	100
	Total:	733.27	2650
	(82% of a 30x30 mm SOAC chip)		

3.3. SOAC Design v.s. COTS Design

The eye-brain machine design based on system-on-a-chip (SOAC) technology has many advantages over the design based on the commercial-off-the-shelf (COTS) components. These SOAC-based design advantages over the COTS-based design are summarized as follows:
(A). Higher On-Chip System Integration: System level integration facilitates innovative parallel processing architecture and ultra-fast data transfer structure while increasing system robustness. (B). 10x Lower Power Dissipation: SOAC EBM consumes much less power than COTS EBM. COTS EBM systems tend to be inherently power hungry. This is because needing a lot of drivers for inter-board and inter-chip connections'capacitive load, needing external bus interface circuit to achieve acceptable data transfer efficiencies. A COTS system typically requires 100 Watts but an SOAC system requires 10 W for the same data throughput.
(C). Speed & Performance: The SOAC EBM can operate at faster frame rates and greater processing speeds. On-chip circuits facilitates shorter interconnections, fewer contacts and drivers, faster devices, more efficient processing architecture and data transfer structure.
(D). Affordability and Compactness: SOAC EBM has on-chip integrated circuitry to reduce cost and size (10x improvement in system miniaturization). The SOAC EBM uses Very Large Scale Integration to

incorporate all necessary vision functions onto one chip. This CMOS integration allows for a very compact system, which increases reliability and reduces cost.

The following sections describe technical details of each on-chip system of the proposed smart vision system.

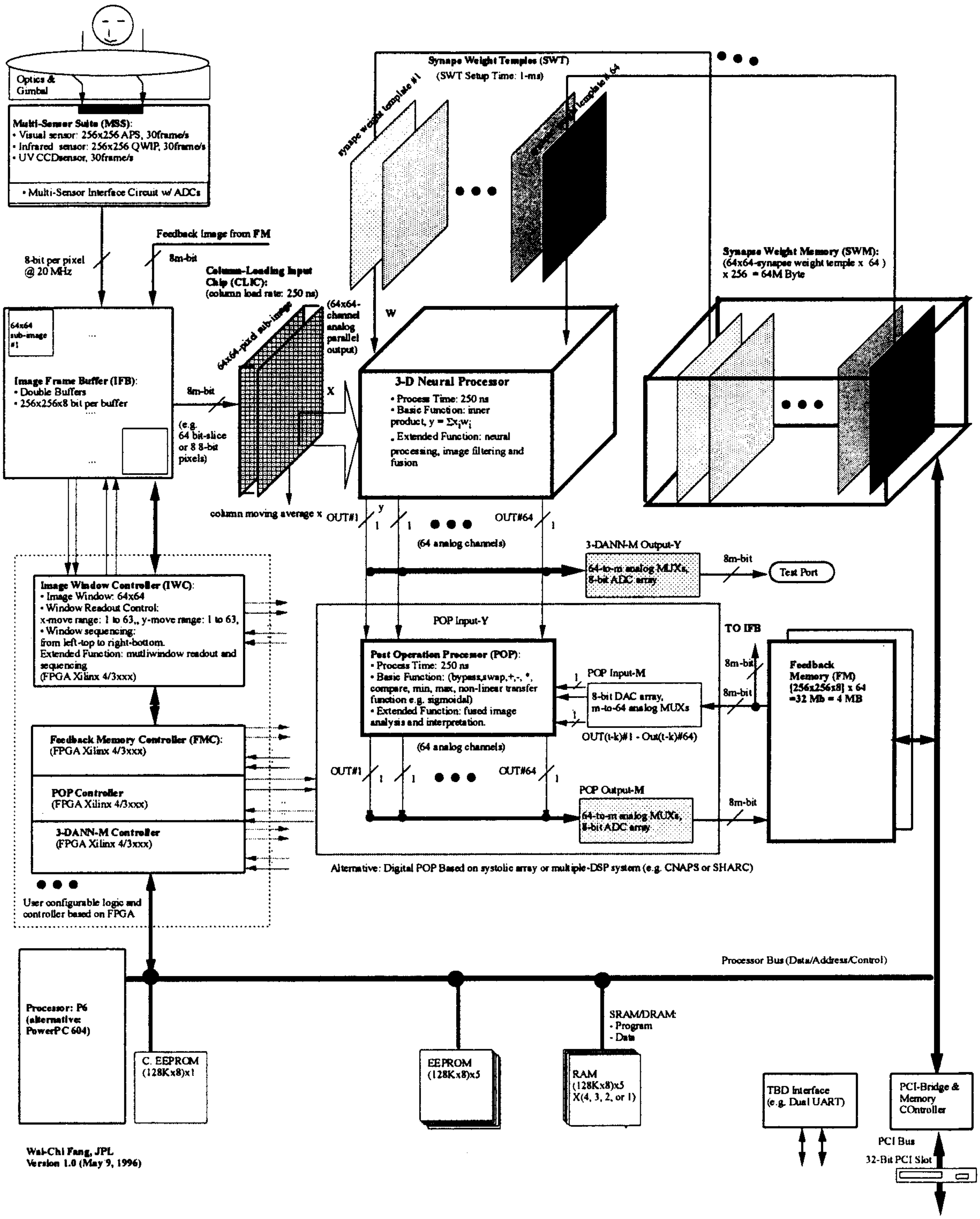

Figure 3.1. An example system design of the eye-brain machine for the VIGILANTE.

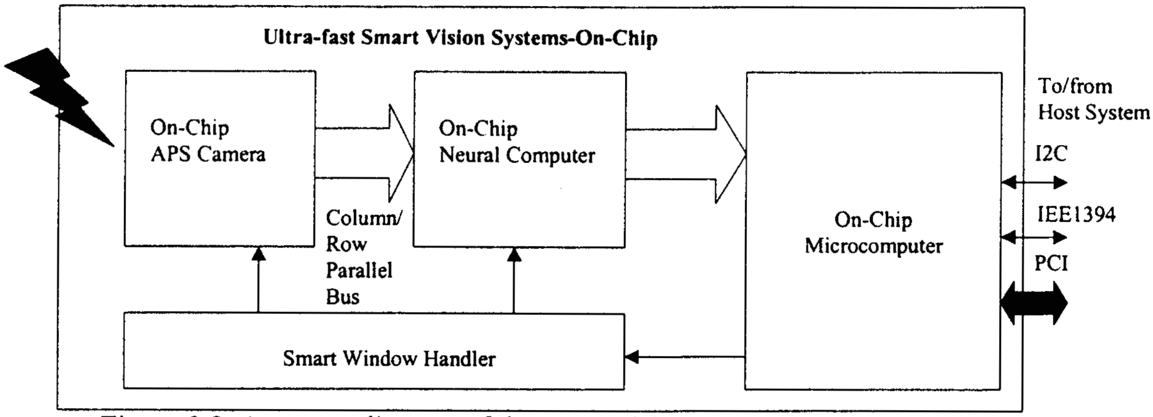

Figure 3.2. A system diagram of the smart vision system.

4. APS Camera-on-a-chip

4.1. CMOS Active Pixel Sensor Technology

High performance CMOS active pixel sensor technology has been developed by NASA's Jet Propulsion Laboratory for electronic image capture [6]. CMOS APS advantages over the CCD are summarized as follows:

(A). 100x Lower Power Dissipation: CMOS APS consumes much less power than CCD systems. A CCD system typically requires 2-5 Watts (digital output) but an APS system requires 20-50 mW for the same pixel throughput.

(B). Affordability and Compactness: CMOS APS has on-chip integrated circuitry to reduce cost and size (10x improvement in system miniaturization).

(C). Windowed Read-out: CMOS APS allows random access to pixel regions of interest. This column and row addressability allows for windows of interest readout, which can be, utilized for machine vision applications needing image compression, motion detection or target tracking.

(D). Image Quality: High quality CMOS APS images have no artifacts, smear or blooming.

(E). Speed & Performance: CMOS APS sensors can operate at faster frame rates.

(F). On-Chip Processing: CMOS APS has on-chip circuits to realize smart functions.

In the APS, both the photodetector and readout amplifier are part of each pixel. This allows the integrated charge to be converted into a voltage in the pixel, which can then be read out over X-Y wires instead of using a charge domain shift register as in CCDs. This column and row addressability allows for window of interest readout (windowing). This windowed read-out feature provides added flexibility in machine vision applications needing image compression, motion detection or target tracking.

4.2. On-Chip APS Camera Design

The 1024x1024 APS is used as the optical sensing array and integrated with the neural processor to build the smart vision system for high definition vision applications. A functional design of the proposed on-chip APS camera is shown in Figure 4.1. These building blocks include an APS active sensor, a smart window handler, a front-end image processor (e.g. data compression) and an image frame memory. The on-chip camera performs image capture at a rate up to 30 frames per second. The on-chip camera can feed the neural processor with input data in a format of mxm sub-window which shift in x rows and y columns basis through the whole image, where x and y are integer ranged from 0 to n-1. Where m and n are dimensions of the camera and the neural computer, respectively. The smart image window handler is designed for the row/column-parallel interface between the APS chip and the neural processor to achieve an ultra-fast frame time.

4.3. A Prototype APS Camera-on-a-chip

A low power 1024x1024 CMOS APS (operate from a +3.3 V supply) using 0.5 μm n-well process was designed and characterized at JPL. Testing results show that the large format APS with small feature size (10-micron pixel pitch) is capable of excellent imaging performance.

A block diagram and chip layout of the 1Kx1K APS prototype chip are shown in Figure 4.2. It contains a 1024x1024 photodiode or photogate pixel array and 1024 parallel 10-bit singles-slope ADC. The analog outputs are VS_OUT (signal) and VR_OUT (reset), and the digital outputs are D_out0 to D_out9. The analog and digital readout chains are separated by the pixel array. Each imager can be operated in analog or digital readout mode. A full image of the 1Kx1K photodiode CMOS APS from the analog output is shown in Figure 4.3.

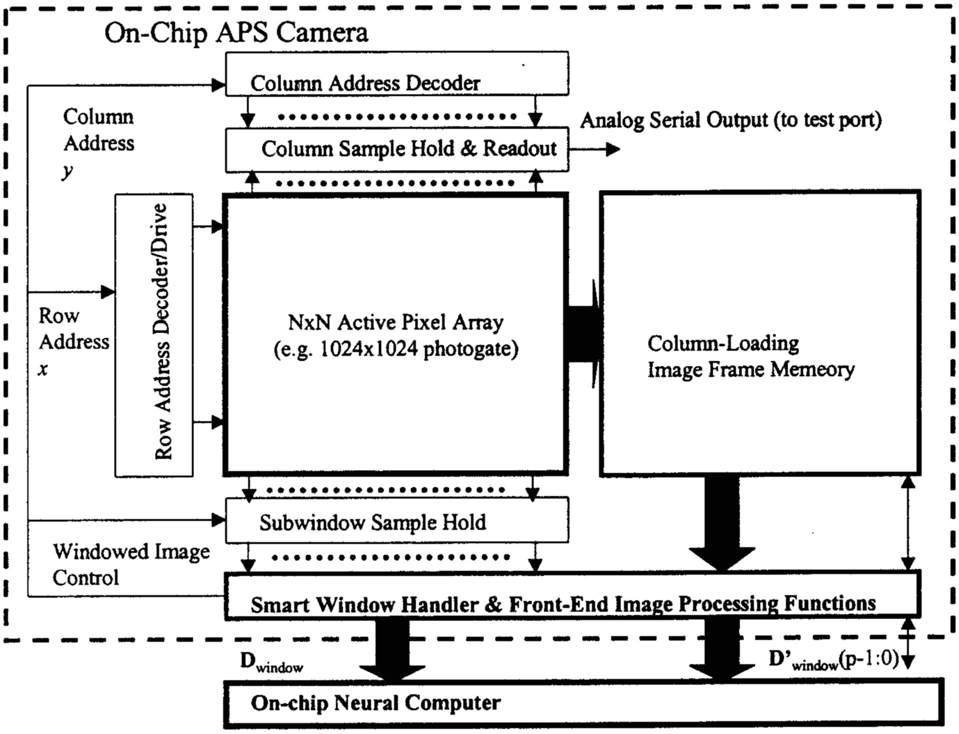

Figure 4.1. A Functional Design of the On-Chip Camera.

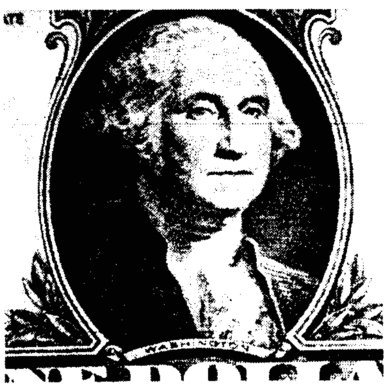

Figure 4.2. Full Image of 1Kx1K Photodiode CMOS APS from the Analog Output.

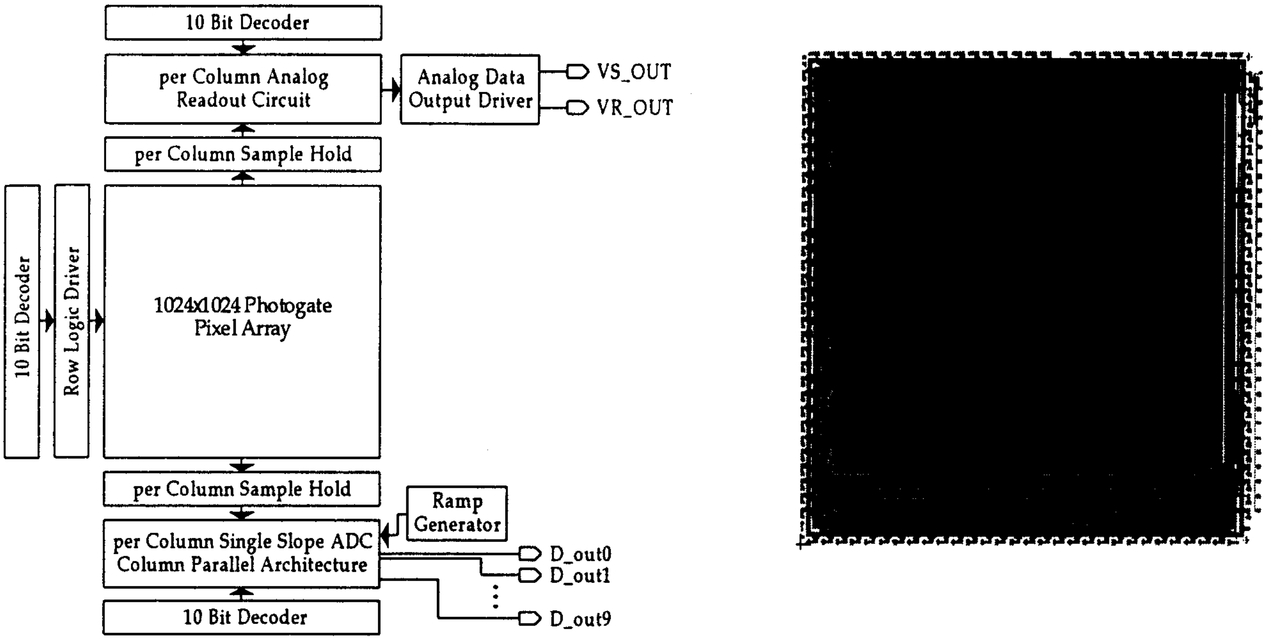

Figure 4.2. (a) Block diagram of 1Kx1K CMOS APS chip. (b) Layout of 1Kx1K CMOS APS.

5. On-Chip Neural Computer

The building blocks of the proposed on-chip neural computer are shown in Figure 5.1. These building blocks include a programmable neural processor, a learning and post operation coprocessor, a programmable synaptic weight memory and an activity memory. The on-chip neural computer can be used as a vision processor at neighborhood of the on-chip camera to perform various vision functions at very high speed. The neural computer performs the feature synthetic image generation, image fusion and fused image analysis. Incorporating the neural computer into the proposed vision system offers orders-of-magnitude computing performance enhancements for real-time vision tasks.

The programmable neural processor is based on optimization cellular neural network (OCNN) [9]. The OCNN is an improved version of the Cellular Neural Networks (CNN) [7]. The CNN has been proved to be universal as the Turing machine [8]. The OCNN keeps this highly desired programmability and become a versatile vision processor. The operation for different tasks depends primarily on the coefficients of the templates and the procedure to apply them. A template includes the information for synapse weights, threshold values, and boundary conditions. The learning of the templates is by defining an object function using semi- or non-parametric methods to iteratively update the weights. Many OCNN functions have been verified via system simulation. These functions include noise filtering, isolated pixel elimination, hole filling, morphological operations, image enhancement, edge detection, connected component detection, feature extraction, motion detection, motion estimation, motion compensation, object counting, size estimation, path tracking, collision avoidance, minimal and maximal detection, etc.

As shown in Figure 5.2, the OCNN is a multi-dimensional array of mainly identical cells, which are dynamic systems with continuous state variables and locally connected with their local cells within a finite radius. The significant design features of the OCNN are described briefly below.

(A) Optimal Solutions of Energy Function:
Under the mild condition [7], a CNN autonomously finds a stable solution for which the Lyapunov function of the network is locally minimized. To improve the local minimized energy function of the

basic CNN, the annealing capability is included to accommodate the applications in which the optimal solutions of energy function are needed. Hardware annealing [9] is a highly efficient method of finding optimal solutions for cellular neural networks.

(B) Multiple Layers with Embedded Maximum Evolution Functions:

In the original CNN every pixel is represented by one neuron. In the OCNN every pixel can be represented by multiple neurons which form a hyperneuron and execute the maximum evolution function for various profile selections or the multi-sensor data synergy.

(C) Digitally Programmable Synapse Weights:

To improve the fixed synapse weights of the CNN, the programmable synapse weights are designed for the OCNN to accommodate the applications, which require programmable pre-determined operators.

(D) High-speed Parallel External Image I/O:

To improve the data I/O bandwidth of the basic CNN, a 2-D array of optical receivers is integrated with the OCNN to accommodate the applications, which require high speed parallel image I/O.

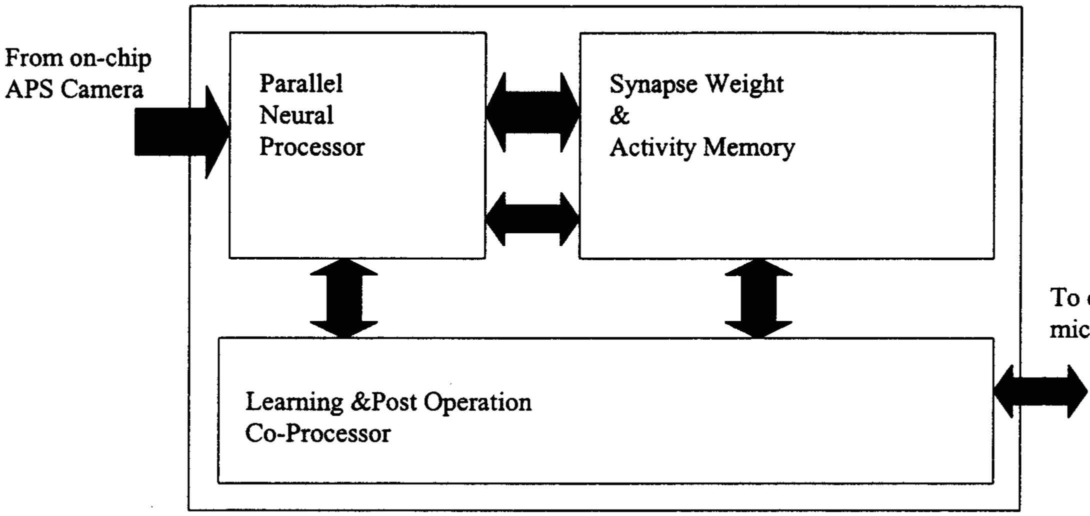

Figure 5.1. A functional diagram of the on-chip neural computer.

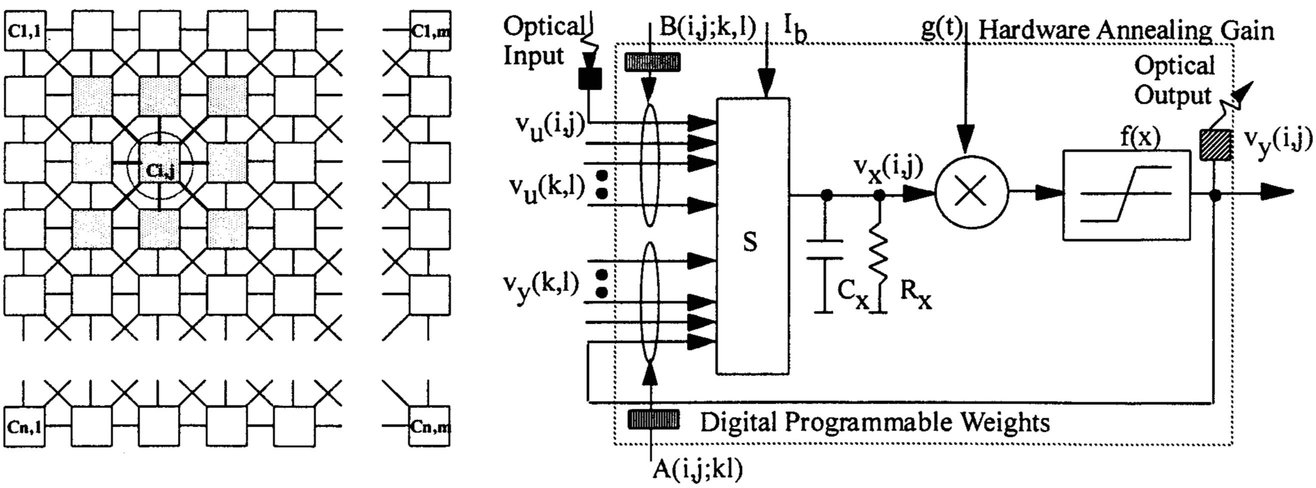

Figure 5.2. (a) An n-by-m OCNN on rectangular grid. The shaded boxes are the neighborhood cells of C(i,j). (b) Functional block diagram of the OCNN neuron C(i,j).

6. On-Chip Microcomputer

The building blocks of the on-chip microcomputer include a microprocessor, a data memory, a program memory, an UART circuit, a clock, a JTAG test support unit, and a multi-bus interface unit. The building blocks of the multi-bus interface unit include an IEEE 1394 interface, an I2C interface, and a PCI bus interface. Host Interface Unit implements the IEEE 1394 bus for high-speed science data and the I2C bus for low speed engineering data. The current version of the IEEE 1394 bus can support data rates of 100 Mbps, 200 Mbps, and 800 Mbps for the cable implementation, and 50 Mbps and 100 Mbps for the backplane implementation. The I2C bus can support a data rate up to 100 Kbps (standard mode) or 400 Kbps (fast mode) and is capable to drive a maximum bus loading of 400 pf. The standard version can address up to 127 nodes (7-bit address) and the extended version can address up to 1023 nodes (10-bit address).

7. Conclusion

A low power smart vision system based on a large format (currently 1Kx1K) active pixel sensor (APS) integrated with a programmable neural processor for fast vision applications is presented. The concept of building a low power smart vision system is demonstrated by a system design, which is composed with an APS sensor, a smart image window handler, and a neural processor. The paper also shows that it is feasible to put the whole smart vision system into a single chip in a standard CMOS technology. This smart vision system on-a-chip can take the combined advantages of the optics and electronics to achieve ultra-high-speed smart sensory information processing and analysis at the focal plane. This highly integrated and ultra-high-speed information processing smart vision system on-a-chip can be used on various NASA scientific missions and other industrial or commercial vision applications.

Acknowledgments

The research described in this paper was performed by the Jet Propulsion Laboratory, California Institute of Technology, and was sponsored by the National Aeronautics and Space Administration. The author would like to thank Dr. Leon Alkalai, Dr. E. Kolawa, Dr. G. Yang, Dr. M. Pain, Dr. S. Udomkesmalee and Professor B. Sheu for valuable technical discussions and supports.

References

[1] D.H. Ballard, C. M. Brown, *Computer Vision*, Prentice-Hall, 1982.

[2] D. Marr, "Vision", Freeman, San Francisco, 1982.

[3] Carver Mead, " Analog VLSI and Neural Systems," Addison-Wesley Publishing Company, 1989.

[4] B. Sheu, J. Choi, Neural Information Processing and VLSI, Kluwer Academic Publishers, 1995.

[5] Suraphol Udomkesmalee, Anil Thakoor, Curtis Padgett, Taher Daud, Wai-Chi Fang, "VIGILANTE: An advanced sensing/processing testbed for ATR applications," SPIE Orlando, FL, April 21-24, 1997.

[6] E.R. Fossum, CMOS Image Sensors: Electronic Camera on a Chip (Plenary paper), Tech. Dig. of IEEE International Electron Devices Meeting, pp. 17-25, Dec. 1995.

[7] L. O. Chua, L. Yang, "Cellular neural networks: Theory", IEEE Trans. on Circuits and Systems, vol. 35, pp. 1257-1272, Oct. 1988.

[8] T. Roska, J. Vandewalle, Eds., Cellular Neural Networks, West Sussex; John Wiley & Sons, 1993.

[9] B. J. Sheu, S. Bang, W.-C, Fang, "VLSI design of cellular neural networks with annealing and optical input capabilities," Proceedings of the IEEE International Symposium of Circuit and System, April 1995.